SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

INSERB INDIA

Indian Institute of Technology Guwahati

7th NATIONAL SYMPOSIUM ON ROTOR DYNAMICS 2024

FAULT CLASSIFICATION IN ROLLING ELEMENT BEARINGS USING MACHINE LEARNING TECHNIQUES

Mirza Samad Beg,
Ravuri Yuvanth,
Prabhu S,
C. Shravankumar
SRMIST, Kattankulathur, India

OVERVIEW

1. Introduction
2. Methodology
3. Description of the work
4. Result and discussion
5. Conclusion

Introduction

•Problem Statement:

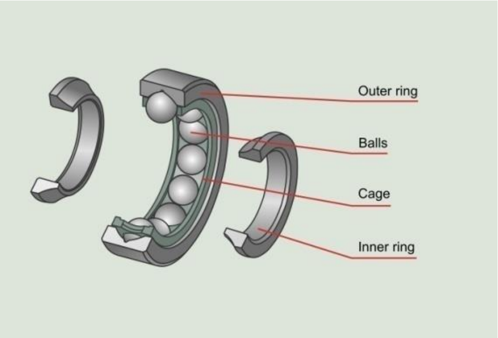
The reliability of rolling element bearings in machinery is critical for reducing downtime and maintenance costs. Fault detection in bearings is a challenge due to the complex behavior of vibrations over time.

•Significance of the Problem:

- Bearings are essential components in rotating machinery, and faults can lead to costly breakdowns.
- Early detection of bearing faults ensures better predictive maintenance, minimizes unplanned downtimes, and prolongs equipment lifespan.

•Applications:

- Manufacturing and industrial automation
- Aerospace and automotive industries
- Energy and heavy machinery sectors



•Objectives of the Work:

- Apply machine learning techniques to classify fault types in rolling element bearings.
- Extract and analyze time-domain and frequency-domain features from vibration signals.
- Investigate different machine learning models for fault diagnosis and predictive maintenance strategies.

Introduction

Fault Type	Description	Frequency Name	Formula	Calculated Value
Inner Race Fault	Damage or defect on the inner race of the bearing, where the rolling elements interact with the inner surface.	BPFI (Ball Pass Frequency Inner): Represents the frequency at which rolling elements pass a specific point on the inner race. This frequency increases with defects on the inner race.	BPFI = $\frac{n}{2}(1+\frac{d}{D}\cos\theta)f_r$	296.90 Hz
Outer Race Fault	Damage or defect on the outer race, the outer surface of the bearing that supports the rolling elements.	BPFO (Ball Pass Frequency Outer): Indicates the frequency at which rolling elements pass a specific point on the outer race. A defect in the outer race produces vibrations at this frequency.	BPFO = $\frac{n}{2}(1-\frac{d}{D}\cos\theta)f_r$	236.43 Hz
Roller Element Fault	Defects such as spalls or cracks on the rolling elements (balls), which roll between the inner and outer races.	BSF (Ball Spin Frequency): Refers to the spin frequency of the rolling elements (balls) themselves. Faults in the balls (e.g., cracks or spalls) cause vibrations at this frequency.	BSF = $\frac{D}{2d}(1-(\frac{d}{D}\cos\theta)^2)f_r$	140.04 Hz

n: Number of rolling elements, **d:** Diameter of rolling elements, **D:** Pitch diameter, **θ:** Contact angle, **fr:** Shaft rotational frequency

Methodology

- Schematic representation of the work.

Data Collection
- IMS Bearing Dataset
- Vibration signals at 20 kHz

Feature Extraction
- **Time-Domain Features**: Mean, RMS, Kurtosis, Skewness, Crest Factor, etc.
- **Frequency-Domain Features**: Peak frequency, amplitude via FFT

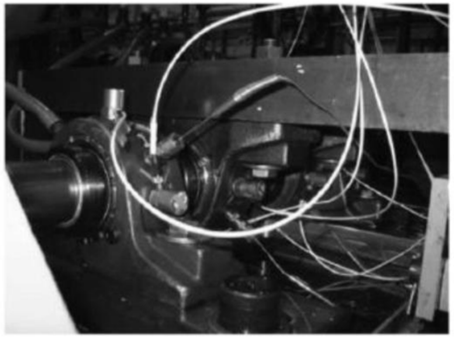
Data Preprocessing
- Normalization
- Splitting into training and testing sets

Model Selection and Training
- Random Forest, Decision Tree, XGBoost, SVM, ANN, Logistic Regression, KNN
- Hyperparameter tuning

Performance Evaluation
- **Metrics:** Accuracy, Precision, Recall, F1 Score, Confusion Matrix

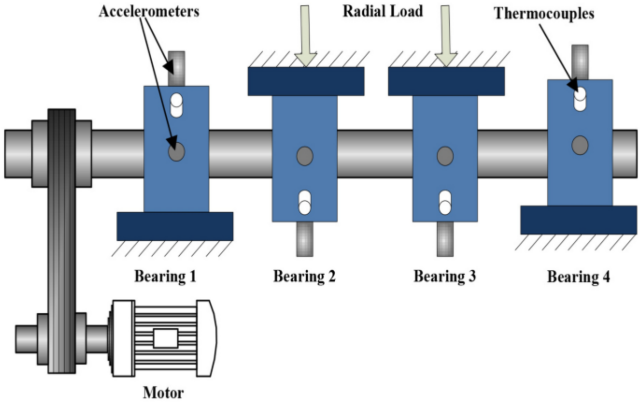
Description of the work

Experimental Setup:



•**Test Rig Setup:**
- •Four **Rexnord ZA-2115 double-row bearings** mounted on a shaft.
- •Shaft driven by an **AC motor** at a constant speed of **2000 RPM**.
- •Radial load of **6000 lbs** applied via a spring mechanism.

•**Instrumentation:**
PCB 353B33 High Sensitivity Quartz ICP Accelerometers mounted on the bearing housing.
- •Sensor arrangement:
  - •Dataset 1: Two accelerometers (x- and y-axes) per bearing.
  - •Datasets 2 & 3: One accelerometer per bearing.



•**Data Collection Details:**
- •Three datasets describing **test-to-failure experiments**:
  - •**Dataset 1:** Inner race and roller element defects.
  - •**Dataset 2:** Outer race defect (Bearing 1).
  - •**Dataset 3:** Outer race defect (Bearing 3).
- •Recording intervals: 10 minutes (5 minutes for initial 43 files in Dataset 1).

•**Failure Progression:**
- •Faults occurred after bearings exceeded their design life of **100 million revolutions**.

Description of the work

Dataset Overview: IMS Bearing Dataset

- **Source:**
  - The dataset is provided by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS).
  - It contains vibration data collected from bearings under varying operating conditions.

- **Dataset Details:**
  - Vibration signals collected using accelerometers placed on the bearings.
  - Sampling frequency: **20 kHz**.
  - Total data points per file: **20,480 points**.

- **Key Features:**
  - Covers three primary types of faults: **Inner race defects, outer race defects, and roller element defects**.
  - Includes healthy bearing data for baseline comparison.

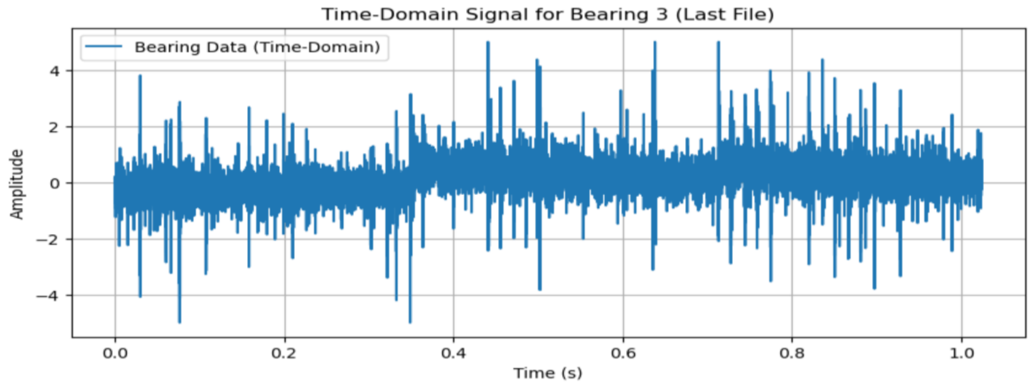
- **Objective:**
  - Analyze vibration signals to detect and classify bearing faults.

Description of the work

Feature Engineering for Fault Diagnosis:

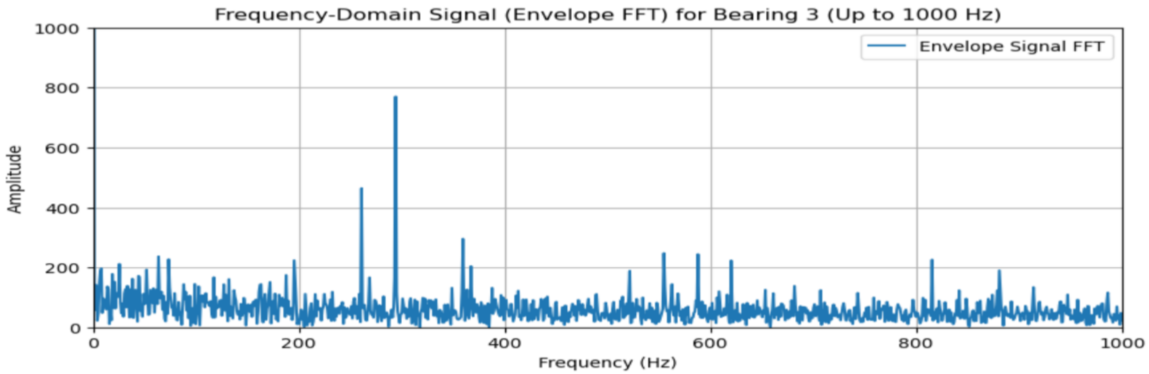
- **Objective of Feature Extraction:**
  - Transform raw vibration signals into meaningful metrics for machine learning models.
  - Capture both **time-domain** and **frequency-domain** characteristics.

- **Time-Domain Features:**
  - **Maximum, Minimum, Mean, Standard Deviation (Std):** Basic statistical metrics.
  - **RMS (Root Mean Square):** Signal power indicator.
  - **Skewness:** Asymmetry of the signal distribution.
  - **Kurtosis:** Sharpness or flatness of the signal peak.
  - **Crest Factor:** Ratio of signal peak to RMS.
  - **Form Factor:** Ratio of RMS to mean value.



Description of the work

- **Frequency-Domain Features:**
  - **FFT (Fast Fourier Transform):** Frequency analysis of the vibration signal.
  - **Envelope Detection:** Highlight modulated frequencies indicating defects.
  - **Top 5 Peak Frequencies and Amplitudes:** Key indicators of fault type and severity.



- **Additional Feature**
  Revolutions:
  - Added the number of shaft revolutions since the start of the experiment as a feature.
  - Calculated using the time index of each vibration snapshot:

$$\text{Revolutions} = \text{Elapsed Time (s)} \times \text{Rotational Speed (RPS)}$$

Description of the work

- **Table of Extracted Features:**

Index	Max	Min	Mean	Std	RMS	Skewness	Kurtosis	Crest Factor	Form Factor	Peak Frequency 1	Peak Frequency 2	Amplitude 1	Amplitude 2	Revolutions	state
1	0.4	-0.496	-0.09081	0.091463	0.128887	0.0343691	0.405106	3.103493	-1.419275	493.16406	14.64844	90.59393	42.0558421	33	Initial
2	0.493	-0.566	-0.0915	0.091734	0.129562	0.0010971	0.445381	3.805125	-1.416038	493.16406	33.20313	100.8939	40.6184416	20033	Initial
3	0.386	-0.601	-0.09392	0.091815	0.131339	0.0530678	0.389381	2.938962	-1.398465	493.16406	14.64844	97.64808	49.8327765	40033	Initial
1049	0.574	-0.833	-0.11453	0.111536	0.159864	0.0011053	0.795298	3.590547	-1.395842	36.132813	33.20313	51.0929	47.596381	20940033	Middle
1050	0.476	-0.725	-0.11459	0.112921	0.160878	0.0026802	0.758125	2.958758	-1.403927	494.14063	1.953125	55.14934	54.3997314	20960033	Middle
1051	0.686	-0.801	-0.11423	0.113336	0.160912	-0.00525	0.807285	4.263196	-1.408684	14.648438	494.1406	82.7756	55.7539361	20980033	Middle
2155	4.902	-5	-0.10755	0.520587	0.531569	-0.083691	11.26935	9.221756	-4.942327	293.94531	359.375	484.43	378.936505	43060033	Inner race
2156	4.316	-4.875	-0.11718	0.459761	0.474447	-0.303903	9.387377	9.096908	-4.048985	293.94531	359.375	453.9675	366.928725	43080033	Inner race
2157	4.998	-5	0.129027	0.579432	0.59361	0.0671893	5.901806	8.419671	4.600673	293.94531	260.7422	770.1537	464.321311	43100033	Inner race

- *In actual dataset Top 5 Peak Frequencies and Amplitudes were used (This table indicated extracted features of bearing 3 in test 1).*

Description of the work

Model Selection and Evaluation :

- **Objective:**
  - Use extracted features to train and evaluate machine learning models for fault classification.

- **Models Used:**
  - **Random Forest (RF):** High accuracy and interpretability.
  - **Decision Tree (DT):** Simple and effective for classification.
  - **XGBoost (XGB):** Gradient boosting for robust performance.
  - **Support Vector Machine (SVM):** Effective for high-dimensional data.
  - **Artificial Neural Network (ANN):** Captures complex patterns in the data.
  - **K-Nearest Neighbors (KNN):** Distance-based classification.
  - **Logistic Regression (LR):** Baseline model for comparison.

- **Key HyperParameters:**
  - **RF:** n_estimators = 100, max_depth = 25, max_features=sqrt
  - **DT:** max_depth = 8, Pruning Parameter (α = 0), max_features = 20
  - **SVM:** Regularization Parameter (C = 100), Decision Boundary Parameter (γ = 0.1)
  - **ANN:** Number of Hidden Layers = 1, Hidden Layer Size = 50, Learning Rate = 0.01, Epochs = 1000.
  - **KNN:** metric = euclidean, n_neighbors = 1, Weights = Uniform.
  - **LR:** Regularization Strength (C = 10), Penalty = L2, Solver = lbfgs
  - **XGB:** max_depth = 9, n_estimators = 50, learning_rate = 0.1, subsample = 1.0.

Result and discussion

Performance Analysis:

•**Model Performance:**

•**Best Model:** Random Forest achieved the highest accuracy (>99%) with minimal misclassification.

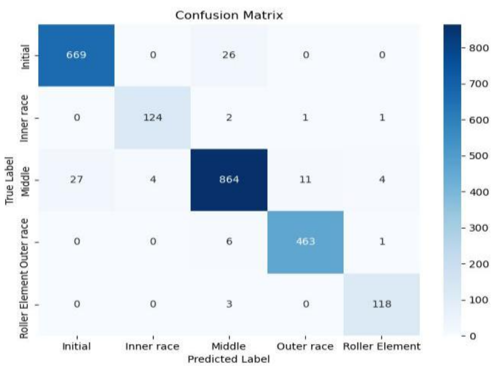
•**Other High-Performing Models:** Decision Tree and XGBoost also performed well, showcasing robust classification ability.

•**Worst Model:** KNN, with an accuracy of 95.11%, struggled due to overlapping class features.

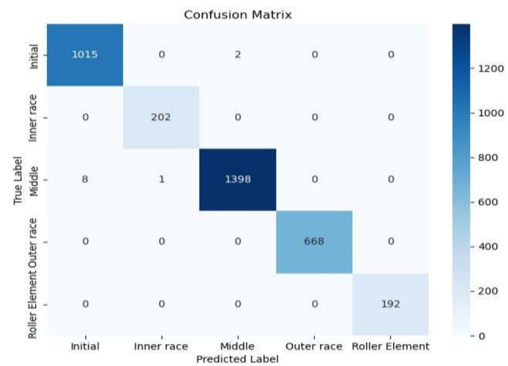
Model	Accuracy (%)
Random Forest (RF)	99.83
Decision Tree (DT)	99.68
XGBoost (XGB)	99.44
Support Vector Machine (SVM)	98.54
Artificial Neural Network (ANN)	96.30
Logistic Regression (LR)	96.24
K-Nearest Neighbors (KNN)	95.11

Result and discussion

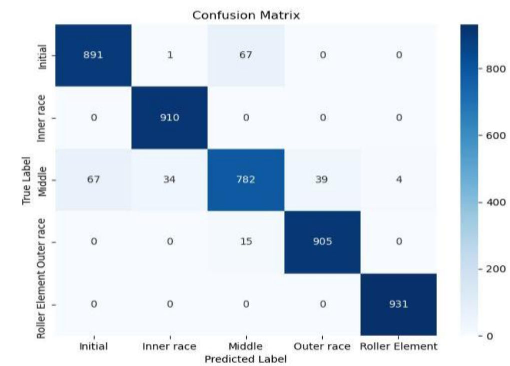
- **Confusion Matrix for the models:**



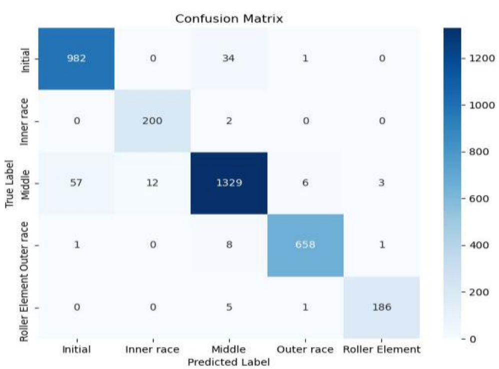
ANN



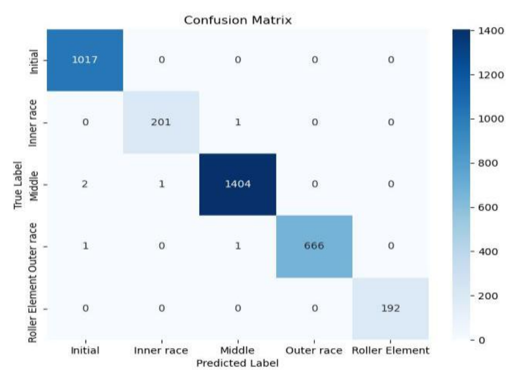
DT



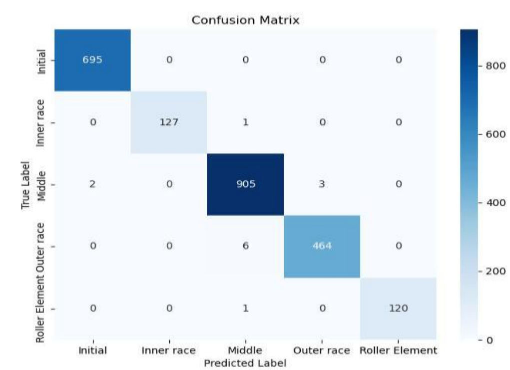
KNN



LR



RF



XGB

Result and discussion

- **Class-Specific Metrics:**
  - **Precision:** RF and XGB achieved perfect precision for most classes, ensuring minimal false positives.
  - **Recall:** RF demonstrated perfect recall for all fault types, meaning no faults were missed.
  - **F1 Score:** RF and DT excelled with near-perfect scores across all classes.
  - KNN struggled with the "Worn Out" class (C3), showing lower recall and F1 scores due to overlapping class features.

Metrics		ML Model						
		ANN	DT	KNN	LR	RF	SVM	XGB
Accuracy		96.30	99.68	95.11	96.24	99.83	98.54	99.44
Precision	C1	0.96	0.99	0.93	0.94	1.00	0.98	1.00
	C2	0.97	1.00	0.96	0.94	1.00	0.99	1.00
	C3	0.96	1.00	0.91	0.96	1.00	0.96	0.99
	C4	0.97	1.00	0.96	0.99	1.00	0.99	0.99
	C5	0.95	1.00	1.00	0.98	1.00	1.00	1.00
Recall	C1	0.96	1.00	0.93	0.97	1.00	0.97	1.00
	C2	0.97	1.00	1.00	0.99	1.00	1.00	0.99
	C3	0.95	0.99	0.84	0.94	1.00	0.96	1.00
	C4	0.99	1.00	0.98	0.99	1.00	0.99	0.99
	C5	0.98	1.00	1.00	0.97	1.00	1.00	0.99
F1 Score	C1	0.96	1.00	0.93	0.95	1.00	0.98	1.00
	C2	0.97	1.00	0.98	0.97	1.00	1.00	1.00
	C3	0.95	1.00	0.87	0.95	1.00	0.96	0.99
	C4	0.98	1.00	0.97	0.99	1.00	0.99	0.99
	C5	0.96	1.00	1.00	0.97	1.00	1.00	1.00

C1 (Healthy), **C2** (Inner Race), **C3** (Worn Out), **C4** (Outer Race), and **C5** (Roller Element)

Conclusion

Key Takeaways and Future Work

- **Summary of Findings:**
  - Machine learning techniques effectively classified bearing faults with high accuracy.
  - **Random Forest** emerged as the best model, achieving accuracy >99% with robust fault classification.
  - **Feature extraction** (statistical and frequency-domain) significantly influenced model performance.

- **Implications:**
  - Enables **early detection of bearing faults**, reducing downtime and maintenance costs.
  - Improves the reliability and safety of rotating machinery in industries like manufacturing and aerospace.

- **Limitations:**
  - Dataset was limited to controlled experimental conditions; further validation is needed for real-world applications.
  - Performance under noisy or incomplete data remains unexplored.

- **Future Work:**
  - Integrate real-time monitoring and fault detection systems.
  - Explore deep learning techniques for automated feature extraction.
  - Extend analysis to larger datasets with diverse fault scenarios.

INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

References

•Gunerkar, R. S., A. K. Jalan, and S. U. Belgamwar. 2019. "Fault Diagnosis of Rolling Element Bearing Based on Artificial Neural Network." Journal of Mechanical Science and Technology 33 (2): 505–511. https://doi.org/10.1007/s12206-019-0103-x.

•Agrawal, Pavan, and Pratesh Jayaswal. 2020. "Diagnosis and Classifications of Bearing Faults Using Artificial Neural Network and Support Vector Machine." *Journal of The Institution of Engineers (India): Series C* 101 (1): 61–72. https://doi.org/10.1007/s40032-019-00519-9.

•A. Meserkhani, S.M. Jafari, A. Rahi, Experimental comparison of acoustic emission sensors in the detection of outer race defect of angular contact ball bearings by artificial neural network, Measurement,Volume 168,2021, 108198, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2020.108198.

•Udmale, Sandeep S., Sanjay Kumar Singh, and Sunil G. Bhirud. 2019. "A Bearing Data Analysis Based on Kurtogram and Deep Learning Sequence Models." *Measurement: Journal of the International Measurement Confederation* 145 (October):665–77. https://doi.org/10.1016/j.measurement.2019.05.039.

•Han, Tian, Zhi Xin Tian, Zhongjun Yin, and Andy C.C. Tan. 2020. "Bearing Fault Identification Based on Convolutional Neural Network by Different Input Modes." *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 42 (9). https://doi.org/10.1007/s40430-020-02561-6.

Thank You

Q & A ?